

# Recommendation-based Keyword Search over Relational Databases

Haithem Ghorbel\*, Nouha Othman\* and Rim Faiz\*\*\*,

\*Université de Tunis, Institut Supérieur de Gestion de Tunis, LARODEC, Tunisia  
gho.haithem,othmannouha@gmail.com

\*\*\*Université de Carthage IHEC Carthage, LARODEC, Tunisia  
rim.faiz@ihec.rnu.tn

**Abstract.** Recently, there has been a burgeoning interest in keyword search in relational databases owing to its ease of use. Although extensive research has been lately done within this context, most of this research not only requires a prior access to data which severely restricts their applicability if this condition is not verified, but also returns very generic answers. However, providing users with personalized answers has become more than ever necessary due to the overabundance of data which can be annoying for the user. The challenge to return personalized and relevant answers that satisfy users' information needs remains. Inspired by the successful application of the collaborative filtering technique in recommender systems, we propose a novel keyword-based approach to provide users with personalized results based on the hypothesis that only information on the database schema is available.

## 1 Introduction

Over the decades, an explosive amount of structured data has been stored in Relational Databases (RDBs). These latter have been widely used owing to the rich information they provide including relationships between the different entities in the DB. Developing effective query methods for users to easily query huge and complex repositories without the need of technical expertise has become one of the biggest challenges of the database community (Agrawal et al., 2002; Aditya et al., 2002). The emergence of web search engines has made keyword search the most commonly used search technique. The strength of this latter is that it enables users to easily express their information needs by a few keywords without needing to know the database schema or structured query languages. Nevertheless, such a technique requires a prior access to the database content in order to build the indices that will pinpoint the different tuples associated with the keywords at run time (Bergamaschi et al., 2011). This is a considerable shortcoming since it limits its applicability if a prior access to data is not possible. Another significant limitation is that the inter-dependencies among the query keywords were ignored. Actually, the meaning of each keyword in a user's query also depends on the meaning of the others. On the other hand, with the tremendous development of information technology, the amount of data has been growing exponentially. Thus, finding the desired information in a massive database has become a crucial but also a challenging task. Recommendation Systems (RSs) are powerful tools to filter data, providing only what the user is most likely looking for. In this paper, we propose a successful attempt to combine RSs techniques and RDB to overcome the limitation of the keyword search. Our proposed approach aims at returning personalized answers when we have no prior access to the actual data stored in the database. The remainder of the paper is structured as follows: In Section 2, we review the main existing work on querying RDBs. Then, we describe our